Biological Journal of the Linnean Society, 2017, **121**, 894–906. With 3 figures.

Geographical diversification and the effect of model and data inadequacies: the bat diversity gradient as a case study

DAVI MELLO CUNHA CRESCENTE ALVES[1*], JOSÉ ALEXANDRE FELIZOLA DINIZ-FILHO[2] and FABRICIO VILLALOBOS[2,3]

[1]*Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, CEP 74.001-970, Goiânia, Goiás, Brazil*
[2]*Departamento de Ecologia, Universidade Federal de Goiás, CEP 74.001-970, Goiânia, Goiás, Brazil*
[3]*Red de Biología Evolutiva, Instituto de Ecología, A.C., Carretera Antigua a Coatepec 351, El Haya, 91070 Xalapa, Veracruz, Mexico*

Received 6 December 2016; revised 3 March 2017; accepted for publication 17 March 2017

The adequacy of some promising phylogenetic comparative methods to test for trait-dependent diversification has been recently criticized to suffer from inflated Type 1 Error rates (i.e. model inadequacy). Nevertheless, formal tests of this model inadequacy for such models within an explicit geographical context are still missing as well as tests of other types of inadequacies such as those related to geographic and phylogenetic data (i.e. data inadequacies). Here, we take advantage of the striking geographic diversity gradient exhibited by bats to explicitly test whether inferences derived from the 'geographic-state speciation–extinction' model (GeoSSE) are biased by model and data inadequacies. We used uncertainty, sensitivity and simulation analyses to show that GeoSSE is sensitive to data inadequacies, being more affected by geographical than phylogenetic inadequacies. Moreover, as previously suggested, the GeoSSE model suffers from inflated Type 1 Error rates. Our results indicate that the GeoSSE model is not reliable for inferring the relative roles of evolutionary processes in driving the bat latitudinal diversity gradient. We argue that uncertainty, sensitivity and simulation analyses should be conducted in all comparative studies that associate species traits and diversification processes to understand diversity gradients.

ADDITIONAL KEYWORDS: character – commission error – macroevolution – phylogenetic uncertainty – species richness – SSE models.

INTRODUCTION

The global species richness of mammals presents the ubiquitous latitudinal diversity gradient (LDG) with a decrease in species numbers from the tropics to the poles (Willig, Kaufman & Stevens, 2003). Although most mammalian orders present such species richness gradient, bats are the main taxon determining the LDG of the whole mammalian class (Kaufman, 1995; Buckley *et al.*, 2010). Therefore, explaining the LDG for bats may not only help to understand the causes driving the mammalian LDG but those of diversity gradients in general since such causes are likely to operate in other taxa as well (Willig *et al.*, 2003; Buckley *et al.*, 2010; Jablonski *et al.*, 2017). Such explanation requires the explicit consideration of the macroevolutionary processes that directly change species numbers: diversification, which is the balance between speciation and extinction, and dispersal (Ricklefs, 2004). Indeed, different evolutionary hypotheses regarding such processes have been proposed to explain large-scale diversity gradients (Mittelbach *et al.*, 2007; Brown, 2014). Thanks to the increasing availability of time-calibrated phylogenies and phylogenetic comparative methods, it is now possible to estimate the rates of macroevolutionary processes and thus discriminate among such evolutionary hypotheses (Pyron & Burbrink, 2013; Morlon, 2014).

For instance, the two main hypotheses advanced to explain the mammalian LDG are the tropical niche

---

*Corresponding author. E-mail: davimello22@gmail.com conservatism (TNC; Wiens & Donoghue, 2004) and out-of-the-tropics (OTT; Jablonski, Roy & Valentine, 2006) hypotheses (Buckley *et al.*, 2010; Rolland *et al.*, 2014). TNC posits that most clades originated in the tropics, occupying it longer and rarely dispersing out of it, thus accumulating more species in that region without implying differences on macroevolutionary rates between tropical and extratropical regions (Wiens & Donoghue, 2004). Whereas OTT also posits a tropical origin of clades but with higher speciation and dispersal and lower extinction rates in the tropics than in extratropical regions (Jablonski *et al.*, 2006). For bats, TNC has been favoured with studies supporting its predictions on their richness gradient (e.g. higher richness of early-diverged species in the tropics and strong positive temperature–richness relationship; Stevens, 2006, 2011; Buckley *et al.*, 2010). However, a recent study considering all mammals contrasted these two hypotheses and found more support for OTT in most orders, including bats (Rolland *et al.*, 2014). Although findings were mostly similar among mammalian orders, some showed contrasting results altogether (e.g. Carnivora) or differences depending on model specifications (e.g. support vs. no support for OTT in Chiroptera) (Rolland *et al.*, 2014). In fact, contrasting results arising from different model specifications may be related to inherent assumptions and biases of phylogenetic comparative methods (Cooper, Thomas & FitzJohn, 2016).

Despite initial excitement on phylogenetic comparative methods that model macroevolutionary processes (the so-called 'diversification models'; Morlon, 2014), several of these methods have been recently criticized and even deemed unreliable (Maddison & FitzJohn, 2015; Rabosky & Goldberg, 2015; Cooper *et al.*, 2016; Moore *et al.*, 2016). For example, the ground breaking model proposed by Maddison, Midford & Otto (2007) that relates the state of a two-state discrete species trait to speciation and extinction rates ['binary-state speciation–extinction' model (BiSSE)] has been shown to be highly sensitive to model violations such as pseudoreplication (Maddison & FitzJohn, 2015) and spurious correlations between a focal trait and diversification rates (Rabosky & Goldberg, 2015). This is particularly important for studies on geographic diversity gradients since the geographical extension of such model ['geographic-state speciation–extinction' model (GeoSSE); Goldberg, Lancaster & Ree, 2011], in which macroevolutionary (speciation, extinction and dispersal) rates are associated to particular regions (e.g. tropics vs. temperate), may suffer from the same issues as the BiSSE model. The GeoSSE model has been widely applied to assess the influence of macroevolutionary processes in determining the geographic diversity gradients of different taxa (Jansson, Rodríguez-Castañeda & Harding, 2013), from plants (Goldberg *et al.*, 2011; Staggemeier *et al.*, 2015), birds (Pulido-Santacruz & Weir, 2016) and reptiles (Pyron, 2014) to the above-mentioned study of mammals (Rolland *et al.*, 2014). Hence, a critical open question is to what extent model assumptions and biases affect inferences from the GeoSSE model.

The most important problem of these state-dependent speciation–extinction models (xxSSE; FitzJohn, 2012) is potentially inferring an association between a species trait and macroevolutionary rates when in fact none exists (model inadequacy; Rabosky & Goldberg, 2015). This problem could induce inflated Type I Error rates, rendering xxSSE models inadequate for testing evolutionary hypotheses (Rabosky & Goldberg, 2015). In addition, uncertainties related to the data used to fit such models (e.g. trait measurements and species' phylogenetic relationships) can also affect the performance of xxSSE models. Regarding the GeoSSE model, such data inadequacies (Fig. 1) can come from the designation of species membership to particular geographic regions, which is based both on defining such regions and on identifying the region(s) within which each species occurs (Goldberg *et al.*, 2011; Fig. 1a, b, respectively) as well as from the phylogenetic uncertainties such as polytomies (Fig. 1c).

Here, we evaluate the influence of model and data inadequacies on inferences derived from the GeoSSE model by means of uncertainty, sensitivity and simulation analyses. We focus on large-scale species richness gradients and the discrimination among evolutionary hypotheses explaining such gradients. For this, we used the striking LDG exhibited by bats. As previously stated, bats are widely used to understand geographic diversity gradients given their high diversity (~1300 species), broad occupation of almost all terrestrial habitats and the considerable amount of phylogenetic and geographic information available for this group (Jones *et al.*, 2002; Willig *et al.*, 2003; Buckley *et al.*, 2010; Peixoto *et al.*, 2014; Shi & Rabosky, 2015). Moreover, several studies had already applied diversification models to understand bats' evolutionary history (Jones *et al.*, 2005; Yu, Wu & Yang, 2014; Shi & Rabosky, 2015), including the GeoSSE model (Rolland *et al.*, 2014), which allows comparison with our findings. Specific results for bats under the GeoSSE model showed contrasting results between constrained and unconstrained dispersal parameters (Rolland *et al.*, 2014) with the former supporting the OTT hypothesis, whereas the latter supporting a reverse trend with lower tropical diversification compared to extratropical regions and higher dispersal from these into the tropics. We show that, at least for bats, such findings and thus supporting a particular evolutionary hypothesis using the GeoSSE model can be heavily dependent on geographic, and less so on phylogenetic,

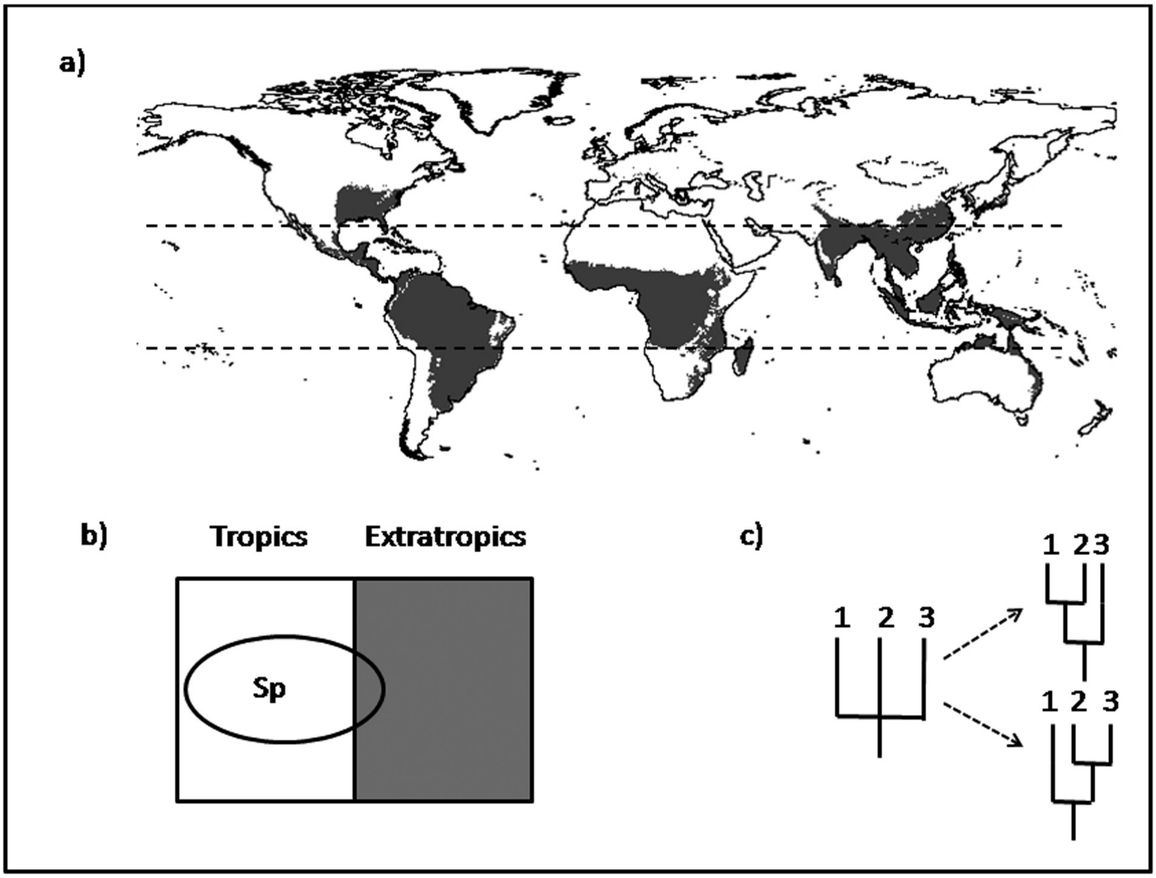

Figure 1. Diagram representing the three types of data inadequacies that could affect inferences from the GeoSSE model for geographic gradients of biodiversity. (a) Represents two regionalization schemes to categorize the globe into tropical and extratropical regions, one based on latitude (superior dotted line: 23.4°N, inferior dotted line: −23.4°S) and another based on environmental productivity (dark grey = tropics; white = extratropics). (b) Represents the commission error on the geographic range of a hypothetical species (elipse) that is endemic to the tropics (white rectangle) but might be considered transtropical because 5% of its range mistakenly overlaps the extratropics (gray rectangle). (c) Represents the generation of two dichotomic phylogenetic trees owning to the 'break' of the polytomy of the original phylogenetic tree.

uncertainties of the input data in addition to suggested model inadequacies.

METHODS

THE GEOSSE APPROACH AND DATA INADEQUACIES

The GeoSSE model is a trait-dependent diversification model, based on the likelihood-based framework of Maddison *et al.* (2007), that uses reconstructed phylogenies of extant species and in which speciation and extinction rates are influenced by the values of a particular species trait (Morlon, 2014). Contrary to the original BiSSE model, where such macroevolutionary rates are tied to the binary trait state (e.g. phenotypic or life history), in GeoSSE the trait is the geographic location of species and thus macroevolutionary rates are tied to both geographic regions where species occur. In addition, a species can occupy one of the two regions or occupy both regions. Finally, state transitions in GeoSSE represent range dynamics of dispersal (expansion) and local extirpation (contraction) (Goldberg *et al.*, 2011). Therefore, GeoSSE requires phylogenetic and distributional information of species as input data.

Geographic data inadequacies

Geographic data for GeoSSE come directly from the distribution of species, either from point occurrences (e.g. Goldberg *et al.*, 2011) or range maps (e.g. Rolland *et al.*, 2014). Such information is then used to define the membership of species to particular regions. For example, in the context of the LDG, species need to be assigned to a particular region such as tropical (*t*, occurring exclusively within the Tropics), extratropical (*e*, occurring exclusively within regions out of the Tropics) or transtropical (*te*, occurring over both regions). Defining the geographic membership of species to such regions requires two steps: (1) determine which regions across the globe are tropical and which are extratropical and (2) identify the region(s) within which each species occurs. The first step can be done in different ways, where most GeoSSE studies have used latitude to categorize the globe into tropical and extratropical regions (e.g. ±23.4°; Fig. 1a; Jansson *et al.*, 2013; Rolland *et al.*, 2014). However, this latitude-based regionalization may be too coarse to define tropical and extratropical regions. For instance, some regions characterized by low average temperature and precipitation are environmentally similar to extratropical regions (<–23.4° or >23.4°) but are considered tropical under a regionalization strictly based on latitude. This is the case of the Mediterranean Forests, Woodlands and Scrubs ecoregion that occurs on high elevations of the Central Andes in South America (Olson *et al.*, 2001). Similarly, some regions characterized by high average temperature and precipitation are environmentally similar to tropical regions (>–23.4° and <23.4°) but are considered extratropical under a regionalization strictly based on latitude, such as the Flooded Grasslands and Savannas ecoregion (i.e. Everglades) in southeast North America (Olson *et al.*, 2001). At large spatial scales, such as those used for studying LDGs, range maps are usually the norm for geographic data (Hurlbert & Jetz, 2007). Consequently, the second step in defining species membership to a given geographic trait state for GeoSSE – that is identifying the region(s) within which each species occurs – is generally done by overlapping species range maps onto tropical and extratropical regions (e.g. Rolland *et al.*, 2014). Range maps represent a coarse model of species geographic distributions and are generated either by experts, which based on their knowledge of species determine the regions where the species can occur, or by simply tracing a minimum convex polygon around the most disperse occurrence points known for each species (IUCN, 2001). On the one hand, range maps are more efficient to reduce omission errors – incorrectly inferring that a species does not occur in a given region – than other geographic data such as point occurrences or species distribution models (Rondinini *et al.*, 2006). On the other hand, range maps unfortunately tend to increase commission errors – incorrectly inferring that a species occurs in a given region (Fig. 1b; Rondinini *et al.*, 2006; Hurlbert & Jetz, 2007; La Sorte & Hawkins, 2007). Under the GeoSSE framework, commission errors could have drastic consequences on the definition of species membership to a given region. For example, if a species is actually adapted to tropical regions but 1% of its range is mistakenly considered to be within extratropical regions, this species will be categorized as transtropical. Hence, if there is a high number of species that are actually adapted to a particular region but their geographic distribution presents commission errors on the tropical-extratropical transition, the amount of transtropical species could be considerably inflated.

Phylogenetic data inadequacy

Phylogenetic data for GeoSSE, and xxSSE model in general, relies on time-calibrated dichotomic resolved phylogenies. However, such phylogenies can suffer from several uncertainties from topology to temporal calibration (Diniz-Filho *et al.*, 2013). For instance, in molecular phylogenies, topological uncertainties such as polytomies can be introduced by the posterior inclusion of species with no molecular data (Rangel *et al.*, 2015). One way to handle such phylogenetic uncertainty on diversification analyses is to break polytomies using, for instance, birth–death models (Kuhn, Mooers & Thomas, 2011) and then use the resultant set of phylogenies in the analyses (Rolland *et al.*, 2014). Nevertheless, there is no consensus on whether this procedure bias the inferences made by phylogenetic comparative methods when estimating diversification rates (Kuhn *et al.*, 2011; Rabosky, 2015). For example, it has been suggested (but not tested) that breaking polytomies under a birth–death model might bias inferences made by trait-dependent diversification models given that the inclusion of unsampled species is not random with respect to the trait distribution among the tips of the phylogeny (Rabosky, 2015).

GEOGRAPHIC AND PHYLOGENETIC DATA OF BATS

We obtained information on bat species phylogenetic relationships from a widely used species-level and time-calibrated supertree of mammals provided by Bininda-Emonds *et al.* (2007) and based on Jones *et al.* (2002) for bats. This supertree was updated by Fritz, Bininda-Emonds & Purvis (2009) and contains phylogenetic information for 1054 bat species. Information on the geographic distribution of bats was obtained from range maps available on the IUCN online database (IUCN, 2014) and, when necessary, we complemented these with information from Wilson & Reeder (2005). There were 1140 species with available geographic data (range maps), and we used this information to determine species membership to tropical, extratropical or transtropical regions across the globe. We adopted the taxonomic classification of Wilson & Reeder (2005) and corrected for all synonyms.

HANDLING DATA INADEQUACIES

Handling geographic data inadequacies

The first step before applying the GeoSSE model to bat data was to determine species membership to one (tropical, extratropical) or both regions (transtropical) across the globe. To deal with the problem of a regionalization scheme solely based on latitude, we generated two alternative regionalizations (i.e. type of traits; hereafter, TRAIT; Fig. 1a). The first TRAIT was the traditional one based on latitude (hereafter, GEO-TRAIT). For GEO-TRAIT, we overlaid the range maps of all bat species on a global map and identified whether a species occurred within the tropical region (i.e. >−23.4° and <23.4°), extratropical region (i.e. <−23.4° or >23.4°) or within both regions. Following Jansson *et al.* (2013), we coded species as '*t*' (tropical), '*e*' (extratropical) and '*te*' (transtropical).

The second TRAIT was based on an environmental variable (hereafter, ENV-TRAIT; Fig. 1a). We assumed productivity (i.e. the amount of biomass in an ecosystem) as the main environmental variable characterizing tropical and extratropical regions, given that high productivity regions are commonly associated with tropical biomes, whereas low productivity regions are usually associated with extratropical biomes (Hawkins *et al.*, 2003). We used actual evapo-transpiration (hereafter, AET) as a proxy for productivity. From a set of productivity-like variables such as net primary productivity and the bioclimatic variables derived from temperature and precipitation, AET was the only one that satisfactorily separated high productive regions, such as tropical humid forests, from low productive regions, such as deserts or high mountain tops (maps not shown). We used a raster file with AET values on a resolution of 0.25° as provided by UNEP (2014).

AET is a continuous variable; thus, we had to transform it into a binary variable to delineate the two regions: tropical and extratropical. We used a *k*-means clustering method (Legendre & Legendre, 2012) to divide the globe, using the raster cells, into two regions. Thisk-means method applies an algorithm to cluster the cells into two groups and identify the cluster that minimizes the difference between the cells within each group (Legendre & Legendre, 2012). We randomly clustered the cells 20 times and used 1000 iterations for each clustering to relocate the cells between the two groups and calculate the within-group residual sum of squares. Finally, to determine species membership to each region based on ENV-TRAIT, we overlaid the range maps of all bat species with a global map of productivity and identified whether a species occurred in a tropical (i.e. high productivity), extratropical (i.e. low productivity) or in both regions. Species were coded in the same way as for the GEO-TRAIT (*t*, *e* and *te*).

Once regions were defined, the second step was to determine the membership of species to each or both regions. To do so, we used range maps that, as mentioned above, may contain commission errors that could inflate the number of transtropical species (Fig. 1b). To deal with this problem, we generated range thresholds (hereafter, RANGE), which consisted in the percentage of species range area (km^2) overlapping the extratropical region. We established 21 range thresholds, ranging from 0 to 20%. Thus, at one end of the spectrum, if we assumed a RANGE of 0%, all species with 0% of their range area overlapping the extratropical region were considered tropical and all species with 100% were considered extratropical. Accordingly, species with 1–99% of their range area overlapping the extratropical region were considered transtropical. At the other end of the spectrum, if we assumed a RANGE of 20%, all species with ≤20% of their range area overlapping with the extratropical region were considered tropical and all species ≥ 80% were considered extratropical. In the same vein, species with 21–79% of their range area overlapping the extratropical region were considered as transtropical.

We assumed range thresholds of 0–20% because there is empirical evidence suggesting a minimum threshold of 20% to realistically consider sympatry between species range (see Price *et al.*, 2014 and references therein), which could also be used to represent commission errors between species range and its occurrence within a region. Considering higher range thresholds could inflate the number of endemic species, instead of inflating the number of transtropical species (i.e. the opposite to the situation explained above).

Handling phylogenetic data inadequacy

The original phylogenetic supertree of mammals, from which we obtained the phylogenetic relationships among bats, presents several polytomies generated by inserting species with no genetic data on the phylogeny (Jones *et al.*, 2005; Bininda-Emonds *et al.*, 2007). To address the uncertainty generated by the breaking of such polytomies in our analyses, we used 100 dichotomic pseudoposterior phylogenies provided by Kuhn *et al.* (2011). These trees were built using a birth–death model to randomly insert the missing species on the phylogeny, given all the taxonomic information available to minimize the error associated with this species input. In addition, we used a maximum clade credibility tree analysis to identify the bat phylogeny with the most common topology among the 100 pseudoposterior phylogenies (MCC; Drummond *et al.*, 2012).

GeoSSE APPLICATION TO BATS LDG

We applied the GeoSSE model (Goldberg *et al.*, 2011) to estimate speciation, extinction and dispersal rates of bats from tropical and extratropical regions. Each parameter is associated to a region and these can be speciation (S*t*, S*e* or S*te*), extinction (X*t* or X*e*) and dispersal (D*t* or D*e*). Because GeoSSE only allows one evolutionary event (speciation, extinction or dispersal) at each time interval, transtropical extinction (X*te*) is not considered given that this would represent two events of extinction (X*t* and X*e*). For the same reason, GeoSSE does not consider transtropical dispersal (D*te-t* or D*te-e*) but simply range contractions determined by extinction events [e.g. a transtropical species becoming a tropical species owing to its extinction from the extratropical region (X*e*); for more details, see Goldberg *et al.*, 2011]. GeoSSE requires two inputs: a phylogeny and a trait vector representing species membership to the regions. In our case, we used several phylogenies to consider phylogenetic uncertainty and trait vectors representing our two regionalization schemes as well as different range thresholds. We randomly selected a sample of 10 phylogenies (PHYs) out of the 100 resolved phylogenies obtained from Kuhn *et al.* (2011; see explanation above). Given that we had two regionalizations (TRAIT) and 21 range thresholds (RANGE), we worked with a total of 42 trait vectors. Therefore, we had 420 combinations of phylogenies and trait vectors (10 phylogenies × 42 trait vectors; hereafter, phy-geo data). Considering all these phy-geo data, we ran the unconstrained GeoSSE model (i.e. all parameters free to vary) for each data. Given that the considered phylogenies were incomplete with regard to the total number of recognized bat species, we used a correction function for the GeoSSE model, as provided in the *diversitree* package, to associate the missing species with the available trait states (FitzJohn, Maddison & Otto, 2009).

UNCERTAINTY ANALYSIS

To quantify the level of uncertainty in GeoSSE parameters associated with geographic and phylogenetic data inadequacies, we partitioned the total parameters variance across all three potential sources of uncertainty represented by our different data: TRAIT, RANGE and PHY (for a similar approach, see Diniz-Filho *et al.*, 2009; Rangel *et al.*, 2015). To do so, we used a PERMANOVA (Anderson, 2001) to understand how much of the variation in GeoSSE parameters (i.e. St, Se, Ste, Xt, Xe, Dt, De) was associated with TRAIT (2 levels: ENV or GEO) and RANGE (21 levels: 0–20% thresholds; Fig. 2). We used PHY as replicates; hence, the residuals of the PERMANOVA were associated with the differences across phylogenies generated by the polytomy resolution. We evaluated each factor separately (TRAIT or RANGE) as well as their interaction (TRAIT × RANGE). To avoid replicate dependency among treatments, we randomly sampled for each treatment 10 PHYs out of the 100 available PHYs, leaving a total of 420 replicates (i.e. 10 replicates per treatment). We used the mean square of each factor to identify how much each of them contributed to the total parameters variation (Gotelli & Ellison, 2004).

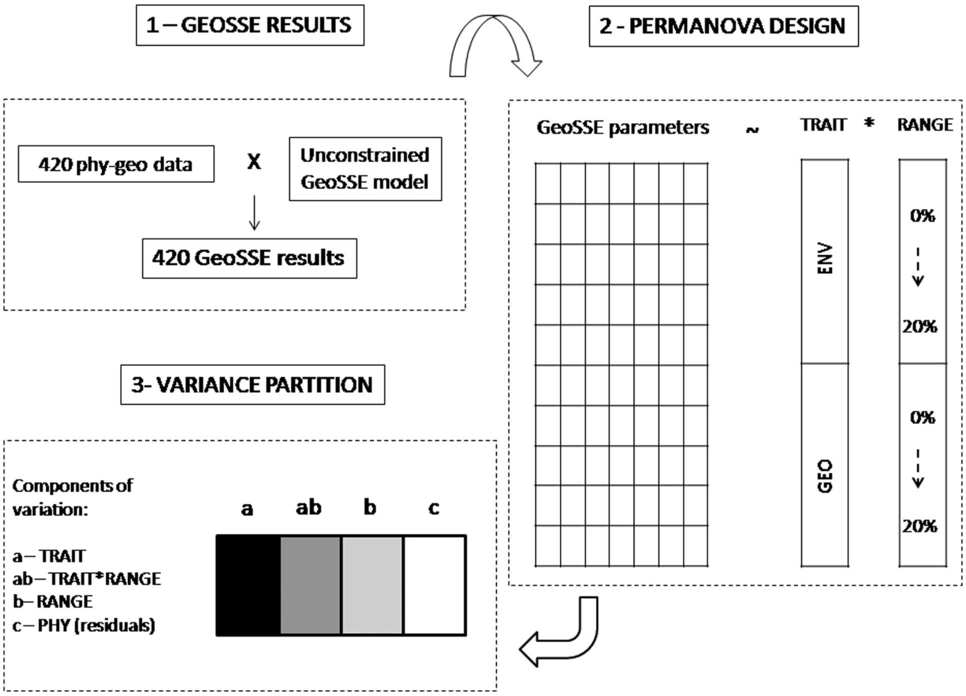

Figure 2. Variance partition of resulting GeoSSE parameters. There were 420 GeoSSE results for the unconstrained model (one for each combination of geographic and phylogenetic data set). ENV, environmental; GEO, geographical; PHY, phylogenetic component or residuals; RANGE, range threshold; TRAIT, regionalization type; *, interactions.

SENSITIVITY ANALYSIS

To determine the effect of geographic data inadequacies on the inferences derived from the GeoSSE model, we evaluated the support of GeoSSE results for a particular hypothesis explaining the LDG (a brief explanation of each hypothesis is given in Supporting Information, Appendix 1). To do so, we followed two steps: (1) express the most common evolutionary hypotheses for the LDG in terms of GeoSSE macroevolutionary parameters (Table 1) and, then, (2) associate GeoSSE results (parameters) with the corresponding hypotheses. Based on this hypotheses-parameter association, sensitivity to geographic data inadequacies was identified as the variation of such association as a function of TRAIT and RANGE factors. Evolutionary hypotheses to explain the LDG can be easily expressed in terms of GeoSSE parameters (Table 1). We considered nine such hypotheses whose underlying mechanisms could be explicitly associated with the macroevolutionary parameters for each region obtained with the GeoSSE model (for an example of such association, see Roy & Goldberg, 2007). Based on this hypotheses-parameters association, we only considered the GeoSSE results for the MCC tree. Thus, we were able to associate each of our 42 GeoSSE results (one for each TRAIT and RANGE) with each of the nine evolutionary hypotheses considered. Note that these hypotheses do not contemplate all parameter combinations, thus some GeoSSE results could not be associated with any hypothesis.

SIMULATION ANALYSIS

To test model inadequacy of GeoSSE, we first simulated a 'null hypothesis' scenario of no association between trait and macroevolutionary parameters (Rabosky & Goldberg, 2015). For the 'null hypothesis' scenario, we simulated 100 phylogenies with the same number of species as our empirical bat phylogeny under a pure-birth process. Next, on each phylogeny, we simulated a neutral trait evolving under a continuous-time discrete-state Markov process. To account for the effect of these neutral traits evolving at different rates, we simulated these traits at four rates (q): 0.05, 0.1, 1 and 10 (Rabosky & Goldberg, 2015). We only used the simulated traits that contained three states and, to avoid Type 2 Error rates, contained more than 10% of species on each state (Davis, Midford & Maddison, 2013). Then, we reshuffled the trait states across the tips to generate random species values. Thus, we simulated the phylogenies as well as random traits to create a conservative 'null hypothesis' scenario of no association between trait and macroevolutionary parameters (Burin *et al.*, 2016). Later, we fitted two GeoSSE models to each simulation data set: (1) a null model where speciation and extinction were constrained to be equal across character states (St = Se; Xt = Xe), while dispersal was potentially asymmetric (Dt ≠ De) and (2) an alternative model where extinction rates were constrained to be equal across states (Xt = Xe), but speciation and dispersal rates were potentially asymmetric (St ≠ Se; Dt ≠ De). Then, we used a likelihood ratio test with significance level of 0.05 to compare model fits. Because our 'null hypothesis' scenario simulated no association between trait and macroevolutionary parameters, we considered an inflated Type 1 Error rate of the GeoSSE model if the best fit model for the simulated data were the asymmetric parameter model (i.e. the alternative model described above). To account for other less conservative 'null hypothesis' scenarios, we also tested GeoSSE for model inadequacy by simulating random and neutral traits under

Table 1. Association between macroevolutionary parameters of the GeoSSE model with evolutionary hypotheses explaining the latitudinal diversity gradient

Model	Hypotheses	Speciation	Extinction	Dispersal
1	Pure dispersal	St = Se	Xt = Xe	Dt < De
2	Macroevolutionary source-sink	St >Se	Xt = Xe	Dt > De
3	Evolutionary speed	St > Se	Xt = Xe	Dt = De
4	Stability	St = Se	Xt < Xe	Dt ≥ De
5	Out-of-the-tropics	St > Se	Xt < Xe	Dt > De
6[*]	Tropical conservatism	St = Se	Xt = Xe	Dt = De
7[†]	Into-the-tropics 1	St > Se	Xt < Xe	Dt < De
8[‡]	Into-the-tropics 2	St = Se	Xt > Xe	Dt < De
9[§]	Into-the-tropics 3	St < Se	Xt < Xe	Dt < De

D, dispersal; e, extratropical region; S, speciation; t, tropical region; X, extinction.

[*] This hypothesis also assumes that the values of dispersal are low.

[†] See Pyron and Wiens (2013) for amphibians.

[‡] See Rolland *et al.* (2014) for an unconstrained model used for bats (see their supporting information).

[§] See Pyron (2014) for squamates.

bat's empirical phylogenies (Supporting Information, Appendix 2).

We performed all the analyses in R (version 3.2.3) using the following packages: *ape* (Paradis, Claude & Strimmer, 2004), *diversitree* (FitzJohn, 2012), *geiger* (Harmon *et al.*, 2008), *phytools* (Revell, 2012) and *vegan* (Oksanen *et al.*, 2015). R code for assessing Type 1 Error rates is available in Supporting Information, Appendix 3.

RESULTS

Our principal aim was to conduct different sets of analyses to evaluate the influence of data and model inadequacies on the inferences derived from the GeoSSE model for the LDG of bats. Our first analysis evaluated uncertainty as the effect of data inadequacies on the variability of GeoSSE model parameters. Indeed, this uncertainty analysis showed that the variance of GeoSSE results is mostly explained by geographic data inadequacies rather than phylogenetic data inadequacies. The factor that contributed the most to parameters variation was RANGE (40.8%), followed by TRAIT × RANGE (33%), TRAIT (16.1%) and PHY (10.1%).

Our second analysis evaluated the sensitivity of inferences derived from GeoSSE (i.e. support for an evolutionary hypothesis) to geographic data inadequacies. This sensitivity analysis showed that supporting a particular evolutionary hypothesis is greatly affected by geographic data inadequacies (Table 2). Considering the environmental regionalization (ENV-TRAIT), 20 GeoSSE results (95.2%) supported the OTT hypothesis and only one result supported a parameter combination not considered in our stated hypotheses. Conversely, considering GEO-TRAIT, 66.6% of the GeoSSE results supported parameter combinations not considered in our hypothesis (14 results), 14.3% of the results supported the Into-the-Tropics 3 hypothesis (4 results), 9.5% of the results supported the Out-of-the-Tropics hypothesis (2 results) and only one result supported the Into-the-Tropics 2 hypothesis (Table 2).

Finally, our last analysis evaluated whether the GeoSSE model systematically infers an incorrect association between geographic occurrences and diversification rates. This simulation analysis showed that GeoSSE model suffers from inflated Type 1 Error rates (Fig. 3). In fact, the GeoSSE model showed inflated Type 1 Error rates for all rates in which the simulated random traits evolved: 48% for $q = 0.05$; 37% for $q = 0.1$; and 20% for both $q = 1$ and $q = 10$. Therefore,

Table 2. Sensitivity of the GeoSSE model to trait (environmental or geographical) and range thresholds (0–20%) expressed as the support for different evolutionary hypotheses

Range thresholds (%)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Hypotheses										Environmental trait											
Pure dispersal																					
Source-sink																					
Evolutionary speed																					
Stability																					
Out-of-the-tropics	■	■	■		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Conservatism																					
Into-the-tropics 1																					
Into-the-tropics 2																					
Into-the-tropics 3																					
No hypothesis				■																	
										Geographical trait											
Pure dispersal																					
Source-sink																					
Evolutionary speed																					
Stability																					
Out-of-the-tropics	■	■																			
Conservatism																					
Into-the-tropics 1																					
Into-the-tropics 2						■															
Into-the-tropics 3									■	■	■	■									
No hypothesis			■	■	■		■	■					■	■	■	■	■	■	■	■	■

Black areas represent GeoSSE results that support a given hypothesis according to trait and range threshold.

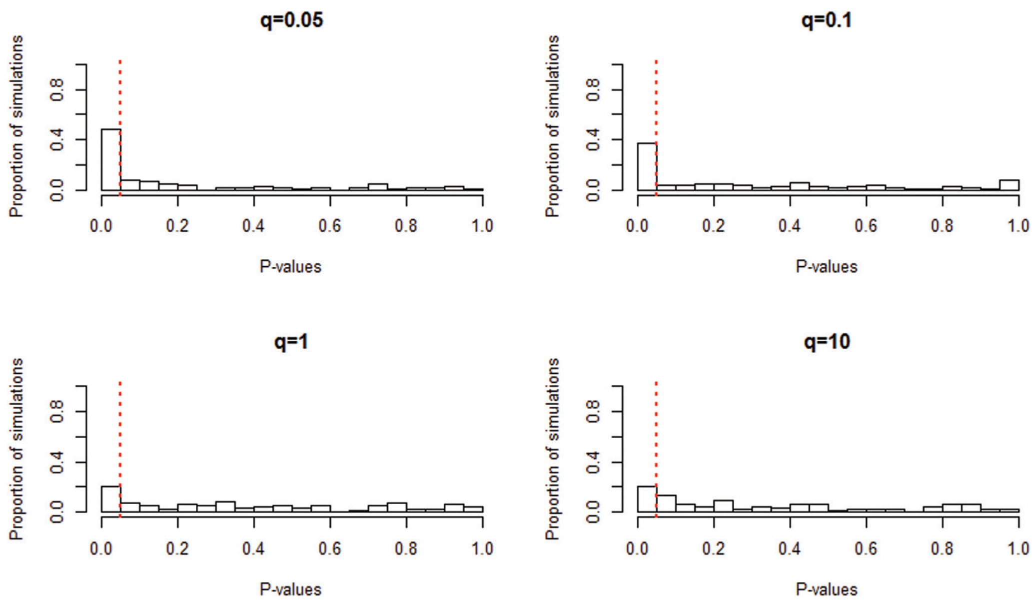

Figure 3. Test of model inadequacy of the GeoSSE model. We simulated 100 phylogenies under a pure-birth process, and for each of them, we simulated a random trait. We repeated this process for traits evolving at four transition rates (q = 0.05, q = 0.1, q = 1 and q = 10). We used the 'likelihood ratio test' to contrast a null and alternative GeoSSE models and evaluate whether GeoSSE incorrectly rejected the null hypothesis of no association between trait and speciation rate. Red dotted lines represent the significance level of 0.05.

even when the disassociation between trait and macroevolutionary rates increased as the trait evolved more rapidly (i.e. higher q rates), the GeoSSE model still presented inflated Type 1 Error rates. Moreover, GeoSSE model also presented inflated Type 1 Error rates for the others 'null hypothesis' scenarios (Supporting Information, Figs 1 and 2 in the Appendix 2).

DISCUSSION

Several hypotheses have been put forward to explain the LDG in terms of ecological and evolutionary processes (Pianka, 1966; Mittelbach *et al*., 2007; Brown, 2014). Compared to ecological processes, evaluating the influence of evolutionary processes has exploded in recent years thanks to the availability of phylogenetic information and comparative methods (Morlon, 2014). However, critical evaluation of such advances, particularly methodological ones, is needed to guarantee the adequacy of data and models used to infer the causes behind diversity patterns (Rabosky & Goldberg, 2015; Cooper *et al*., 2016). We have highlighted the need for such critical evaluation when using a geographical diversification model to support evolutionary hypotheses explaining the LDG. Using the striking LDG exhibited by bats as an example, we showed that the GeoSSE model is not only affected by model inadequacy, as previously suggested for the BiSSE model (Rabosky & Goldberg, 2015), but also by data inadequacies, namely geographic and phylogenetic. Such model and data inadequacies can severely bias our inferences, potentially leading us to support an incorrect evolutionary hypothesis.

The type of data inadequacy that most affected the GeoSSE model was the way in which species membership was assigned to a particular region (RANGE, in our terminology; Fig. 3). This suggests that studies that have used GeoSSE to understand geographical gradients of species richness for groups such as amphibians, mammals and squamates (Pyron & Wiens, 2013; Pyron, 2014; Rolland *et al*., 2014), might have reached biased conclusions (but see Pulido-Santacruz & Weir, 2016 for an exception). For instance, a recent study of mammals (Rolland *et al*., 2014) applied the GeoSSE model under a geographical regionalization (GEO-TRAIT, in our terminology) and no restriction on range overlap (RANGE of 0%, in our terminology) to define species membership to tropical and extratropical regions. Particularly for bats, this study found that both regions had the same speciation rate but different extinction and dispersal rates, with higher extinction in tropical regions and higher dispersal from the extratropics into the tropics (Rolland *et al*., 2014; based on the unconstrained GeoSSE model). In our analysis, using similar geographic and phylogenetic data as well as model specifications (unconstrained dispersal) as Rolland *et al*. (2014) for bats, we were able to support the same explanation (i.e. into-the-tropics 2 hypothesis) only under GEO-TRAIT and RANGE of 5% (Table 2). This finding implies that our interpretation of the evolutionary processes responsible for the bat LDG is highly dependent on geographic data, mainly the level of commission error assumed in the analysis.

The notion that commission errors associated with range maps can mask the understanding of geographic patterns of biodiversity is not new (Hurlbert & Jetz, 2007; La Sorte & Hawkins, 2007). Recently, different studies have explicitly considered this data inadequacy in their analyses by validating, assuming or testing range thresholds. For instance, Tobias *et al.* (2014) established a threshold of 20% of breeding range overlap among ovenbird species to determine whether they were sympatric (i.e. >20%) or allopatric (i.e. < 20%). This threshold was validated by a systematic revision of published species range maps and point occurrences (Tobias *et al.*, 2014). Other authors incorporated commission errors in their analyses by assuming a unique threshold of 25% of latitudinal range overlap with tropical and extratropical regions; for example a species was considered tropical only if more than 75% of its range was tropical (Kerkhoff, Moriarty & Weiser, 2014). And recently, Pulido-Santacruz & Weir (2016) tested RANGEs of 5, 10, 15 and 20% for birds on GeoSSE and found no significant differences across their results. This latter study further suggested that the sensitivity of GeoSSE to RANGE might be clade specific. Therefore, whenever accurate geographic data to determine species membership to a given region is missing, different RANGEs should be considered when applying and interpreting results from the GeoSSE model.

After the range threshold factor (RANGE), the data inadequacy that most affected GeoSSE was the way in which we categorized the globe into tropical and extratropical regions (TRAIT, in our terminology; Table 2). In our results, both regionalizations (TRAITs) were associated with different evolutionary hypotheses explaining the LDG of bats. When considering the environmental regionalization (ENV-TRAIT), results were associated with two hypotheses, whereas using the geographical regionalization (GEO-TRAIT) associated the results with four hypotheses. A possible explanation for such distinct inferences from the choice of regionalization is that the environmental regionalization produces regions that are more fragmented and detailed given that environmental variables such as primary productivity are heterogeneously distributed on geographic space (Fig. 1a). Consequently, a more fragmented regionalization produced, on average, a greater amount of transtropical species (Supporting Information, Appendix 4), which in turn reduced the parameter variance across the different range thresholds (RANGEs). This parameter consistency across RANGEs favoured the support of fewer hypothesis by GeoSSE. Thus, explicitly considering the environment in delineating different regions seems to be a more consistent scheme for inferences derived from diversification analyses.

Phylogenetic data inadequacy had a lower effect on the variation of GeoSSE parameters than geographic data inadequacies. Our results contradict the expectation that polytomy resolution based on birth-death models could bias inferences from trait-dependent diversification models (Rabosky, 2015). Other studies had already shown no significant biases in diversification patterns owing to phylogenetic uncertainty after breaking polytomies (Kuhn *et al.*, 2011; Rolland *et al.*, 2014). For instance, Rolland *et al.* (2014) found that speciation and extinction rates for several mammalian orders were consistent across their pseudoposterior trees. Thus, we believe that phylogenetic uncertainty caused by the polytomy resolution might not considerably affect inferences made by trait-dependent diversification analyses. Even so, these findings may be clade-specific and future studies should be conducted to test the generality of this potentially negligible effect of phylogenetic uncertainty on diversification analyses.

Our findings suggest that the GeoSSE model, as it has been implemented so far, does not provide fully reliable tests of alternative evolutionary hypotheses (Fig. 3). This model showed inflated Type 1 Error rates in our conservative 'null hypothesis' scenario, which presented a high level of disassociation between trait and diversification rates, implying that the GeoSSE model is prone to associate diversification differences with traits that did not cause such difference. Such spurious correlations are also consistent across less conservative 'null hypothesis' scenarios based on empirical bat phylogenies with neutral and random traits (Supporting Information, Appendix 2). Thus, our results reinforce Rabosky & Goldberg's (2015) expectation that SSE models (BiSSE, ClaSSE, GeoSSE, MuSSE, QuaSSE) might be inadequate to test evolutionary hypotheses. Some possible solutions to overcome this model inadequacy is to use a statistical procedure to correlate lineage-specific diversification rates – generated by a trait-independent model – with a biological trait (Rabosky & Huateng, 2015; but see Moore *et al.*, 2016), or to use a trait-dependent model that includes hidden states in the analyses (HiSSE; Beaulieu & O'Meara, 2016). A drawback of the latter solution is that the model was built for traits that present only two known states. Therefore, the HiSSE model cannot be used, in its current form, to understand the LDG because species geographic membership requires three trait states: tropical, extratropical and transtropical.

Aside from data and model inadequacies, other concerns need to be taken into account when using SSE models to understand geographic diversity gradients. One important issue is the linkage between evolutionary hypotheses and GeoSSE results highlighted in this study (Table 1). Some hypotheses have additional components than simply speciation, extinction and dispersal. For instance, the TNC hypothesis posits that the elapsed time also contributes to the species richness pattern ('time-for-speciation' effect), given that the Tropics are older than extratropical regions thus having more time to accumulate species (Wiens & Donoghue, 2004; see Supporting Information, Appendix 1). This effect, however, cannot be explicitly tested with GeoSSE. Therefore, alternative methods to SSE models are still needed to test all of the components posit by evolutionary hypotheses to explain geographic diversity gradients.

CONCLUSIONS

We demonstrated here that the use of a trait-dependent diversification model to understand geographic patterns of biodiversity is highly biased by data and model inadequacies. Geographic data inadequacies related to the definition of tropical and extratropical regions as well as commission errors of species geographic distributions affect inferences made by the geographical SSE model more than phylogenetic uncertainty. Moreover, as expected, the geographical SSE model demonstrated to be inadequate to test evolutionary hypotheses owing to inflated Type 1 Error rates. These evidences illustrate how problematic SSE models can be to understand geographic diversity gradients. We highlight that the use of uncertainty, sensitivity and simulation analyses to evaluate data and model inadequacies should not be restricted only to the use of SSE models within a geographical context, but to all comparative studies that associate phylogenies with causal factors (e.g. biological traits or abiotic variables) to understand biodiversity patterns.

ACKNOWLEDGEMENTS

We are indebted to Lucas Jardim for helpful comments on the manuscript and Luciano F. Sgarbi for optimizing the R codes. We also thank Folmer Bokma, John A. Allen, Sebastian Höhna, Tanja Stadler and two anonymous reviewers for helpful comments on a previous version of this manuscript. DMCCA received a studentship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). JAFD-F has been continuously supported by CNPq productivity grants. FV was supported by a BJT 'Science without Borders' grant from CNPq.

REFERENCES

Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26:** 32–46.

Beaulieu JM, O'Meara BC. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* **65:** 583–601.

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. 2007. The delayed rise of present-day mammals. *Nature* **446:** 507–512.

Brown JH. 2014. Why are there so many species in the tropics? *Journal of Biogeography* **41:** 8–22.

Buckley LB, Davies TJ, Ackerly DD, Kraft NJ, Harrison SP, Anacker BL, Cornell HV, Damschen EI, Grytnes J, Hawkins BA, McCain CM, Stephen PR, Wiens JJ. 2010. Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proceedings of the Royal Society B: Biological Sciences* **277:** 2131–2138.

Burin G, Kissling WD, Guimarães Jr PR, Şekercioğlu ÇH, Quental TB. 2016. Omnivory in birds is a macroevolutionary sink. *Nature Communications* **7:** 1–10.

Cooper N, Thomas GH, FitzJohn RG. 2016. Shedding light on the 'dark side' of phylogenetic comparative methods. *Methods in Ecology and Evolution* **7:** 693–699.

Davis MP, Midford PE, Maddison W. 2013. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evolutionary Biology* **13:** 1–11.

Diniz-Filho JAF, Bini LM, Rangel T, Loyola RD, Hof C, Nogués-Bravo D, Araújo M. 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography* **32:** 897–906.

Diniz-Filho JAF, Loyola RD, Raia P, Mooers AO, Bini LM. 2013. Darwinian shortfalls in biodiversity conservation. *Trends in Ecology & Evolution* **28:** 689–695.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3:** 1084–1092.

FitzJohn RG, Maddison WP, Otto SP. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology* **58:** 595–611.

Fritz SA, Bininda-Emonds OR, Purvis A. 2009. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters* **12:** 538–549.

Goldberg EE, Lancaster LT, Ree RH. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Systematic Biology* **60:** 451–465.

Gotelli NJ, Ellison AM. 2004. *A primer of ecological statistics*. Sunderland: Sinauer.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W. 2008. GEIGER: investigating evolutionary radiations. *Bioinformatics (Oxford, England)* **24:** 129–131.

Hawkins BA, Field R, Cornell HV, Currie DJ, Guégan JF, Kaufman DM, Kerr J, Mittelbach G, Oberdorff T, O'Brien E, Porter E, Turner JR. 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* **84:** 3105–3117.

Hurlbert AH, Jetz W. 2007. Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. *Proceedings of the National Academy of Science of the United States of America* **104:** 13384–13389.

IUCN. 2001. *IUCN red list categories and criteria. Version 3.1*. Gland: IUCN Species Survival Commission.

IUCN. 2014. *International Union for Conservation of Nature and Natural Resources*. Available at: http://www.iucnredlist.org/

Jablonski D, Huang S, Roy K, Valentine JW. 2017. Shaping the latitudinal diversity gradient: new perspectives from a synthesis of paleobiology and biogeography. *The American Naturalist* **189**.

Jablonski D, Roy K, Valentine JW. 2006. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science* **314:** 102–106.

Jansson R, Rodríguez-Castañeda G, Harding LE. 2013. What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. *Evolution* **67:** 1741–1755.

Jones KE, Purvis A, MacLarnon A, Bininda-Emonds OR, Simmons NB. 2002. A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biological Reviews of the Cambridge Philosophical Society* **77:** 223–259.

Jones KE, Bininda-Emonds OR, Gittleman JL. 2005. Bats, clocks, and rocks: diversification patterns in Chiroptera. *Evolution* **59:** 2243–2255.

Kaufman DM. 1995. Diversity of New World mammals: universality of the latitudinal gradients of species and bauplans. *Journal of Mammalogy* **76:** 322–334.

Kerkhoff AJ, Moriarty PE, Weiser MD. 2014. The latitudinal species richness gradient in New World woody angiosperms is consistent with the tropical conservatism hypothesis. *Proceedings of the National Academy of Science of the United States of America* **111:** 8125–8130.

Kuhn TS, Mooers AØ, Thomas GH. 2011. A simple polytomy resolver for dated phylogenies. *Methods in Ecology and Evolution* **2:** 427–436.

La Sorte FA, Hawkins BA. 2007. Range maps and species richness patterns: errors of commission and estimates of uncertainty. *Ecography* **30:** 649–662.

Legendre P, Legendre LF. 2012. *Numerical ecology*. Amsterdam: Elsevier.

Maddison WP, FitzJohn RG. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* **64:** 127–136.

Maddison WP, Midford PE, Otto SP. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* **56:** 701–710.

Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, Harrison SP, Hurlbert AH, Knowlton N, Lessios HA, McCain CM, McCune AR, McDade LA, McPeek MA, Near TJ, Price TD, Ricklefs RE, Roy K, Sax DF, Schluter D, Sobel JM, Turelli M. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters* **10:** 315–331.

Moore BR, Hohna S, May MR, Rannala B, Huelsenbeck JP. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Science of the United States of America* **113:** 9569–9574.

Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecology Letters* **17:** 508–525.

Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H. 2015. *Vegan: community ecology package, R package, Version 2.2-1*. Available at: http://CRAN.R-project.org/package=vegan

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, D'Amico J, Itoua I, Strand H, Morrison J, Loucks C, Alnutt T, Ricketts T, Kura Y, Lamoreux J, Wettengel W, Hedao P, Kassem KR. 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience* **51:** 933–938.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Peixoto FP, Braga PHP, Cianciaruso MV, Diniz-Filho JAF, Brito D. 2014. Global patterns of phylogenetic beta diversity components in bats. *Journal of Biogeography* **41:** 762–772.

Pianka ER. 1966. Latitudinal gradients in species diversity: a review of concepts. *American Naturalist* **100:** 33–46.

Pulido-Santacruz P, Weir JT. 2016. Extinction as a driver of avian latitudinal diversity gradients. *Evolution* **70:** 860–872.

Price TD, Hooper DM, Buchanan CD, Johansson US, Tietze DT, Alström P, Olsson U, Ghosh-Harihar M, Ishtiaq F, Gupta SK, Martens J, Harr B, Singh P, Martens, J. 2014. Niche filling slows the diversification of Himalayan songbirds. *Nature* **509:** 222–225.

Pyron RA. 2014. Temperate extinction in squamate reptiles and the roots of latitudinal diversity gradients. *Global Ecology and Biogeography* **23:** 1126–1134.

Pyron RA, Burbrink FT. 2013. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends in Ecology & Evolution* **28:** 729–736.

Pyron RA, Wiens JJ. 2013. Large-scale phylogenetic analyses reveal the causes of high tropical amphibian diversity. *Proceedings of the Royal Society B: Biological Sciences* **280:** 20131622.

Rabosky DL. 2015. No substitute for real data: a cautionary note on the use of phylogenies from birth–death polytomy resolvers for downstream comparative analyses. *Evolution* **69:** 3207–3216.

Rabosky DL, Goldberg EE. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* **64:** 127–136.

Rabosky DL, Huang H. 2016. A robust semi-parametric test for detecting trait-dependent diversification. *Systematic Biology* **65:** 181–193.

Rangel TF, Colwell RK, Graves GR, Fučíková K, Rahbek C, Diniz-Filho JA. 2015. Phylogenetic uncertainty revisited: implications for ecological analyses. *Evolution* **69:** 1301–1312.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3:** 217–223.

Ricklefs RE. 2004. Cladogenesis and morphological diversification in passerine birds. *Nature* **430:** 338–341.

Rolland J, Condamine FL, Jiguet F, Morlon H. 2014. Faster speciation and reduced extinction in the tropics contribute to the mammalian latitudinal diversity gradient. *PLoS Biology* **12:** 1001775.

Rondinini C, Wilson KA, Boitani L, Grantham H, Possingham HP. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters* **9:** 1136–1145.

Roy K, Goldberg EE. 2007. Origination, extinction, and dispersal: integrative models for understanding present-day diversity gradients. *The American Naturalist* **170** (Suppl 2): S71–S85.

Shi JJ, Rabosky DL. 2015. Speciation dynamics during the global radiation of extant bats. *Evolution* **69:** 1528–1545.

Staggemeier VG, Diniz-Filho JA, Forest F, Lucas E. 2015. Phylogenetic analysis in Myrcia section Aulomyrcia and inferences on plant diversity in the Atlantic rainforest. *Annals of Botany* **115:** 747–761.

Stephens PR, Wiens JJ. 2003. Explaining species richness from continents to communities: the time-for-speciation effect in emydid turtles. *The American Naturalist* **161:** 112–128.

Stevens RD. 2006. Historical processes enhance patterns of diversity along latitudinal gradients. *Proceedings of the Royal Society of London B: Biological Sciences* **273:** 2283–2289.

Stevens RD. 2011. Relative effects of time for speciation and tropical niche conservatism on the latitudinal diversity gradient of phyllostomid bats. *Proceedings of the Royal Society of London B: Biological Sciences* **278:** 2528–2536.

Tobias JA, Cornwallis CK, Derryberry EP, Claramunt S, Brumfield RT, Seddon N. 2014. Species coexistence and the dynamics of phenotypic evolution in adaptive radiation. *Nature* **506:** 359–363.

UNEP. 2014. United Nations Environment Programme. Available at: http://www.unep.org/

Wiens JJ, Donoghue MJ. 2004. Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution* **19:** 639–644.

Willig MR, Kaufman DM, Stevens RD. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics* **34:** 273–309.

Wilson DE, Reeder DM. 2005. *Mammal species of the world: a taxonomic and geographic reference*. Baltimore: JHU Press.

Yu W, Wu Y, Yang G. 2014. Early diversification trend and Asian origin for extent bat lineages. *Journal of Evolutionary Biology* **27:** 2204–2218.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix 1. Evolutionary hypotheses for the latitudinal diversity gradient.
Appendix 2. Other simulation analyses to test GeoSSE for model inadequacy.
Appendix 3. Model inadequacy functions (R scripts).
Appendix 4. Cumulative proportion of species on each trait (i.e. tropical, extratropical or transtropical) according to 21 thresholds (i.e. percentage of 0–20%) of area of range overlap with extratropical biome. First row corresponds to environmental trait and second row corresponds to geographical trait.